OLECULAR BIOLOGY™ • 221
Editor: John M. Walker

n of cDNA Libraries

ods and Protocols

Edited by

Shao-Yao Ying

enome and the potential application of its information to gene chips,
nctional assays of specific gene sequences has become critical in gene
neration of cDNA Libraries: Methods and Protocols, expert researchers
vn proven techniques for generating cDNA/mRNA libraries to identify
quences. A wide variety of techniques is presented for enhancing the
raries, and for confirming the quality of the cDNAs generated. Among
resis, Northern blotting, single-cell microarray analysis, subtractive
cloning, and peptide library generation. Each method includes back-
ions, a list of reagents, operational tips, and notes on instrumentation.
nformation: the definition of a cDNA library, the various types of cDNA
ls for cDNA library generation using either conventional approaches
NA libraries.

al, *Generation of cDNA Libraries: Methods and Protocols* provides
ucible techniques for the generation of the entire range of complete,
d in today's forefront genetic research.

FEATURES

nd

ls from

- **Single-cell microarray analysis**
- **Discussion of cDNA libraries in diagnostics, drug development, and clinical therapy**

CONTENTS

ISBN 1-58829-066-2
90000

Ying

Generation of cDNA Libraries

Methods in Molecular Biology™
VOLUME 221

Generation of cDNA Libraries

Methods and Protocols

Edited by

Shao-Yao Ying

12

mRNA/cDNA Library Construction Using RNA–Polymerase Cycling Reaction (RNA-PCR)

Shi-Lung Lin and Shao-Yao Ying

1. Introduction

Molecular profiling of single-cell gene expression permits the high-definition investigation of intracellular gene activity and physiological status of the cells under certain special conditions, such as pathogenesis *(1,2)*, cancer staging *(3)*, drug treatment, and developmental processes *(4)*. Traditionally, gene transcripts were extracted from lysed cells with phenol–chloroform followed by precipitation, and messenger RNAs (mRNA) were further purified by oligo-(dT)-dextran media *(5)*. However, the tedious procedures of extraction, chromatography, and precipitation could not maintain the completeness of a whole mRNA repertoire, resulting in a significant loss of rare RNA (<10 copies/cell) populations. Such loss could be as much as 30% of the original repertoire. The requirement of bulk tissue samples for a better population coverage was another drawback of the phenol–chloroform extraction methods. A minimum of several thousand cells is needed for an acceptable quality of RNA extraction. Because of tissue heterogeneity, these methods usually provided neither reliable nor reproducible results. Unfortunately, it is impossible to collect adequate amounts of pure or homogeneous samples for these methods because of a tremendous difficulty in sample dissection and RNA preservation, especially the preservation of rare mRNA species.

A breakthrough improvement of mRNA preparation is now based on a in vitro transcription (IVT) reaction, which provides linear amplification of a whole poly(A)$^+$ RNA repertoire up to 2000-fold per cycle from limited numbers of cells *(6,7)*. By incorporating an RNA promoter into cDNA templates, these transcription-based methods amplified nucleotides by RNA polymerization.

From: *Methods in Molecular Biology, vol. 221: Generation of cDNA Libraries: Methods and Protocols*
Edited by: S.-Y. Ying © Humana Press Inc., Totowa, NJ

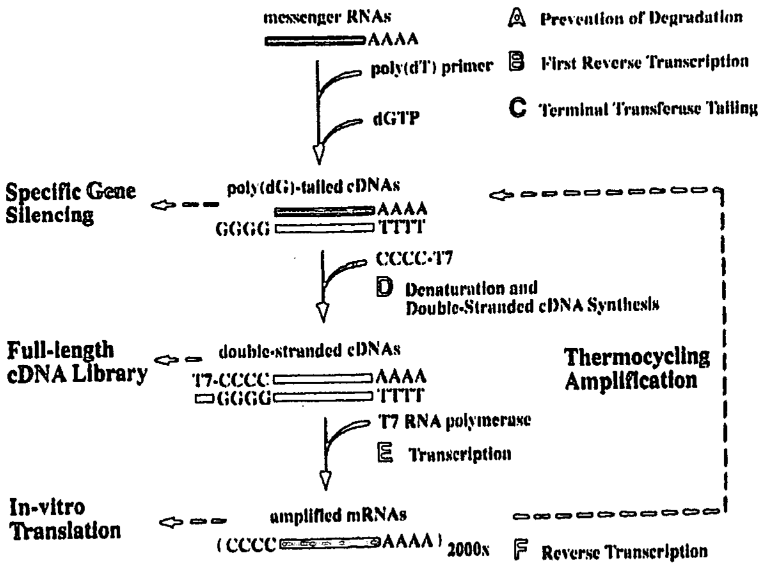


Fig. 1. An illustration of the RNA-PCR thermocycling procedure. The cycling steps D–F can be repeated at least one time for the linear amplification of a mRNA library by in vitro transcription. Advantageously, the reactions of steps A–F can be continuously performed in a reverse transcription and in vitro transcription (RT&T) buffer. The cycling of reverse and in vitro transcription reactions provides more flexibility for the enzymatic synthesis of single-stranded RNAs, RNA–DNA hybrids, and double-stranded DNAs, which are ready for a variety of biochemical applications such as mRNA library preparation for microarray analysis (*see* **Figs.** 3 and 4), probe preparation for specific gene detection *(7)*, full-length gene cloning, in vitro translation for protein synthesis, and gene knockout analysis through a posttranscriptional gene silencing mechanisms *(8)*.

The identification of some useful mRNA markers for certain disease detection has been reported *(2,3)*. Recently, a novel thermocycling procedure, RNA-polymerase cycling reaction (RNA-PCR), further achieved full-length mRNA amplification and successfully displayed cancer-stage-specific geno expression by Northern blot analysis *(3)*. To the best of our knowledge, this is the first procedure that has been tested to generate a full-length mRNA library from as few as 20 tissue cells (2-pg mRNAs) for profiling cancer stages in vivo. In brief, the RNA-PCR (polymerase chain reaction) procedure (*see* **Fig.** 1) is based on (A) prevention of degradation, (B) reverse transcription of mRNAs with poly-(dT) primers, (C) poly(dC)-tailing of the first-strand cDNAs, (D) denaturation and then double-stranding the DNA templates with oligo-(dG)-

T7-promoter primers, (E) in vitro transcription from the promoter region to generate multiple RNA sequences, and (F) repeating steps A–E without step C to achieve the desired poly(A)$^+$ RNA amount for analysis.

This method is capable of generating cell-type-specific poly(A)$^+$ RNA libraries up to 5 kb in the full-length conformation of most mRNAs, and up to 12 kb in a shorter 5′ truncated form of larger mRNA species. A high G-C content RNA, however, tends to be a little shorter than its original size. In general, the good integrity of total cellular mRNAs should appear as a smear between approx 500 bases and 5 kb on an electrophoresis gel and is composed of a median size of around 2 kb *(5)*. It is noteworthy that the full-length conformation at this range actually covers more than 90% of a whole mRNA population in cells. Based on our electrophoresis data, the quality of an RNA-PCR-derived mRNA library has reached the same quality as a smear between 300 and 7.4 kb without ribosomal RNA and genomic DNA contamination on a 1% formaldehyde–agarose gel (*see* **Fig.** 2A). Northern blot analysis of p16, a rare gene usually not shown in phenol–chloroform extracted RNAs, was clearly detected in a RNAPCR-derived library, whereas the signal of GAPDH (a highly abundant gene) was observed in all tested libraries, indicating a better preservation of rare mRNA species by RNA-PCR amplification. Moreover, RB (4.9 kb), β-actin (2.2 kb), and GAPDH (1.7 kb) gene transcripts were all measured in their corrected full-length sizes (**Fig.** 2B), further confirming a potential full-length conformation up to at least 5 kb. The utilization of thermostable reverse transcriptases in our current protocol has improved the full-length potential up to 9 kb and the 5′-end start codon of the resulting mRNA reading frames can be well preserved for further in vitro translation. In addition to the linear amplification of a IVT-based reaction, the RNA-PCR-derived RNAs could, therefore, proportionally represent most of mRNA populations in their original makeup.

To test high-yield and linear amplification, we have routinely generated 30 μg of amplified mRNAs in a 40-μL reaction mixture after three rounds of RNA-PCR amplification from about 20 single cells (approx 1 ng total RNAs). This represents a 1.5×10^7-fold increase based on a comparison between the amount of synthesized poly(A)$^+$ RNAs and that of theoretically presumed mRNAs within a cell (0.1 pg). Even after 10-fold dilution of current enzymatic activities, a more than 20-fold increase of specific mRNA sequences was measured in each cycle of transcriptional amplification (*see* **Fig.** 2C). Such high-yield amplification has been proven to be a linear amplification process, as shown in **Fig.** 2D. Because of the strict proof-reading feature of RNA polymerases, linear amplification is a natural property of transcriptional amplification methods *(7,9)*. Linear amplification maintains the accurate ratio of each expressed gene transcript in an amplified library for representing the

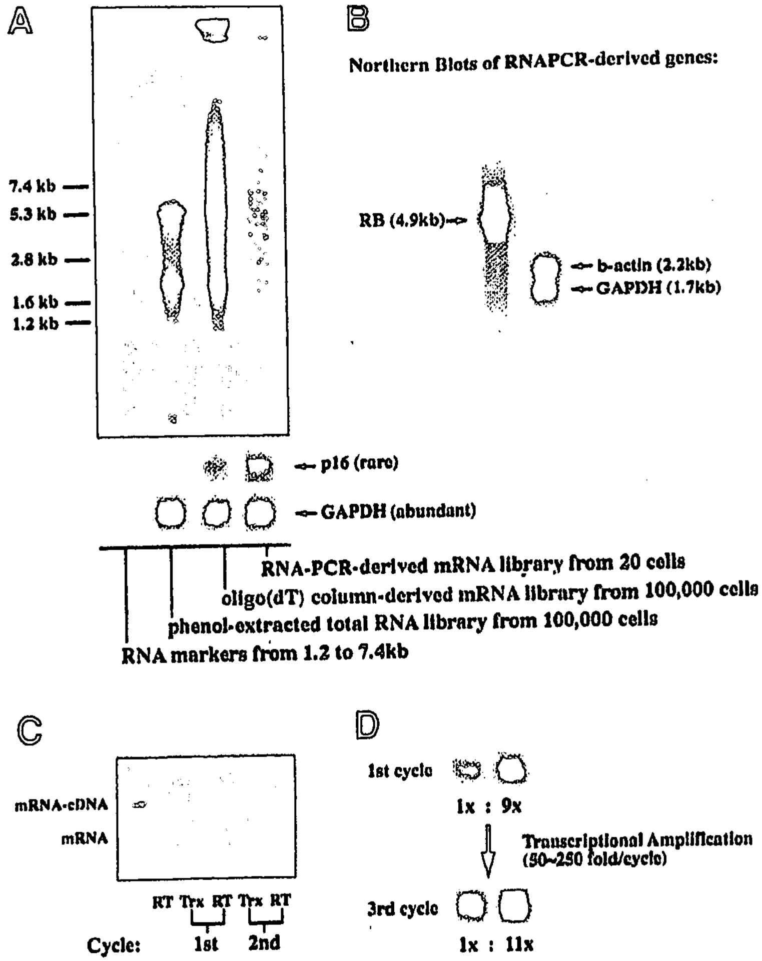
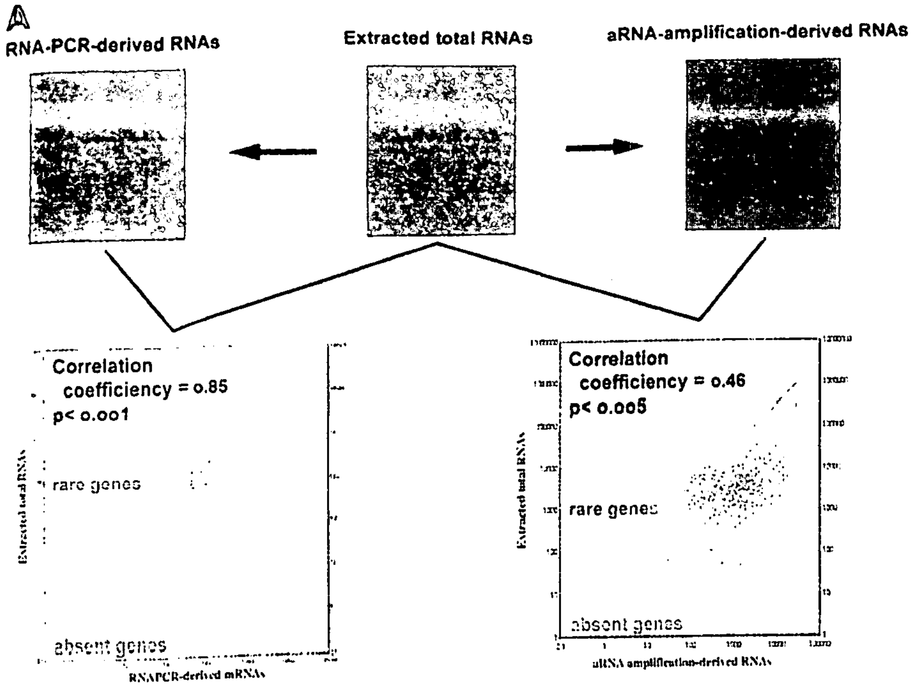


Fig. 2. Analyses of basic RNA-PCR features. (A) Comparison between RNA libraries prepared by phenol–chloroform extraction (lane 2), oligo-(dT) chromatographic column (lane 3), and RNA-PCR (lane 4) fractionated on a 1% formaldehyde–agarose gel, all ranging from 300 to above 7.4 kb based on RNA markers. A uniform smearing pattern of all three products indicates good RNA quality and quantity. p16, a rare and quickly degraded gene transcript, can be clearly identified in the RNA-PCR-derived library but not the others, whereas the abundant GAPDH and β-actin transcript was detected by Northern blots in all three libraries. (B) The amplification level of a specific gene transcript (activin) between two cycles of RNA-PCR has shown a significant 10-fold increase after utilization of a 1/10-fold enzymatic activity (20 U of T7 RNA polymerase). A greater than 250-fold amplification rate has been detected when 200 U of T7 RNA polymerase was applied to a RNA-PCR reaction (3). However, such a tremendous amplification rate cannot be observed by gel electrophoresis without dilution. (C) The ratio of amplified gene products in (D) was analyzed by Northern blotting at two predetermined concentrations (1:9) after two cycles of RNA-PCR amplification. The final ratio (1:11) was considered to closely match the original 1:9 ratio, indicating that the transcriptional amplification is a fairly linear amplification procedure.

original RNA composition. Indeed, the correct ratio composition of a whole mRNA population is critical to warrant the reproducibility and representation of its resulting gene analyses. However, previous methods for RNA preparations could not provide any evidence for linear amplification of rare RNAs *(6,7,9,10)*. To this end, the RNA-PCR-derived RNA library was experimentally tested to provide better lineage, coverage, and representation of RNA amplification for high-density microarray hybridizations, as shown in **Fig. 3**.

Reliable reproducibility and representation of a RNA-PCR-derived mRNA library have been confirmed using microarray analysis. When applied to affymetrix U95A2 gene chips ($n = 3$), both phenol–chloroform extracted and RNA-PCR-amplified RNA libraries displayed about 4200 expressed genes on a total about 12,670 gene chips (33.5 ± 0.3% and 33.2 ± 0.4%, respectively), representing a very similar size of RNA populations. It should be noted that the RNA-PCR-amplified library is amplified from 20 LNCaP cells, whereas the total RNA library is extracted from about 10^6 cells of the same. Among all expressed genes, 17 of them were completely missing in 1 of the libraries, indicating a 0.4% of representation loss that may have occurred during different handling. Less than 2% of the average population (approx 102 genes) was detected to be differentially displayed more than threefold changes, showing a very good mutual representation capacity. From the computing results of scatterplots (**Fig. 3**, left), a highly linear correlation of gene coverage was found in the abundant and moderate mRNA species of both. A more intense signal for rare RNA population was detected in the RNA-PCR-derived mRNAs, indicating a better preservation of most rare species. Although they were not perfectly matched with each other, the above results have demonstrated much more promising compatibility than those from the comparison of the extracted total RNAs to an aRNA library amplified by Eberwine's conventional aRNA amplification method from 20 LNCaP cells (**Fig. 3**, right). It is known that the aRNA amplification has been widely applied to prepare labeled probes for microarray hybridization *(7,11)*. However, because of the utilization of random primers for cycling amplification, the compared aRNA library ($n = 3$) displayed a more skewed and less abundant population containing an average of 2243 expressed genes (approx 17.8%).

Using 8000 gene DNA microarrays provided by the National Cancer Institute (NCI), we have also performed a similar experimental comparison among a standard reference RNA library from NCI, amplified aRNAs generated from 10 ng of the standard, and RNA-PCR-derived RNAs amplified from 10 ng of the standard ($n = 2$). It showed an average 83% linear correlation between the standard and RNA-PCR-derived RNA libraries (*see* **Fig. 4**, middle), whereas only 49% correlation was detected between the standard RNA and aRNA library (**Fig. 4**, right). Based on 0.1 pg mRNA per cell and each cell containing



B

	aRNA amplification	RNA-PCR
mRNA amplification	− (random reverse primers)	+
aRNA generation	+	+ (if using oligo[dT]-T7 primers for resulting mRNA amplification)
Full-length cDNA amplification	− (random reverse primers)	+
Single-cell amplification	+ (only for specific genes)	+ (for a whole library)
Linear amplification	+ (theoretical proposal)	+ (experimental proof)

Fig. 3. The comparison of differential gene expression patterns by microarray analyses. **(A)** Among RNA-PCR-derived poly(A)$^+$ RNAs from 20 cells, aRNAs amplified by Eberwine's method *(6,7)*, and traditional phenol–chloroform extracted total RNAs from 10^6 cells, the results showed a highly compatible population in the abundant and moderate mRNA species of the extracted and the RNA-PCR-derived RNAs (left) but not the aRNAs (right). The rare mRNAs (white zone between two squares) were, however, markedly different among all three groups. The gray shadow area indicated a marginal reading zone of signal detection, showing genes absent in the compared libraries. The 45° green lines parallel to the *x*/*y* slope marked the differential changes from one (near the center) to eight (to both axes) folds. **(B)** A list of distinct features between RNA-PCR and aRNA amplification methods, which may contribute to the different results of microarray analysis.

2% mRNAs, the 10-ng standard of referenced total RNAs is equivalent to the mRNA amount of 200 cells. It is a reasonable cell number for sample collection by a laser capture machine (LCM). For single-cell microarray analysis, this result suggests that RNA-PCR preserves much better mRNA information using fewer single cells than previous methods. However, the current method is not suitable for analysis of bulk RNA sources because of its limitation in terminal deoxynucleotidyl transferase (TdT) tailing procedure. For RNA samples greater than 5 µg, Eberwine's aRNA amplification method is still the best choice *(11)*.

2. Materials

2.1. Generation of First-Strand cDNA Using Reverse Transcription with Poly(dT) Primers

1. Diethyl pyrocarbonate (DEPC) H_2O: Stir double-distilled water with 0.1% DEPC for more than 12 h and then autoclave at 120°C under about 1.2 kgf/cm^2 for 20 min, twice.
2. Poly(dT)$_{24}$ primer: dephosphorylated 5′-dTTTTTTTTTT TTTTTTTTTT TTTT-3′ (100 pmol/µL) (*see* **Note 1**).
3. AMV reverse transcriptase (50 U/µL) and 10X reverse transcription buffer (500 m*M* Tris-HCl [pH 8.5] at 25°C, 80 m*M* $MgCl_2$, 300 m*M* KCl, and 10 m*M* dithiothreitol [DTT]).
4. First reverse transcriptase mix: 7 µL DEPC-treated ddH_2O, 2 µL 10X reverse transcription buffer, 2 µL of 10 m*M* dNTP mix (10 m*M* each of dATP, dGTP, dCTP, and dTTP), 1 µL RNasin (25 U/µL), and 2 µL AMV reverse transcriptase; prepare just before use.
5. Incubation chamber: 65°C, 42°C, and 50°C.
6. Purification spin column: 100 bp cutoff filter (e.g., a Microcon-50 centrifugal filter [Amicon, Beverly, MA], as shown here).

2.2. cDNA Amplification Using Terminal Transferase Tailing with Poly(dC) Oligonucleotides

1. 10X terminal transferase tailing buffer: 500 m*M* Tris-HCl (pH 8.0) at 25°C, 400 m*M* KCl, 80 m*M* $MgCl_2$, and 100 m*M* DTT; prepare fresh.
2. Terminal deoxynucleotidyl transferase (25 U/µL).
3. Terminal transferase reaction mix: 3 µL DEPC-treated ddH_2O, 1 µL of 10 m*M* dNTP, 3 µL of 10X terminal transferase tailing buffer, 1 µL of 10 m*M* dCTP, and 2 µL terminal deoxynucleotidyl transferase; prepare just before use.
4. Incubation mixer: 37°C, 100*g* vortex for 30 s between every 5-min interval.
5. Incubation chamber: 37°C and 94°C.

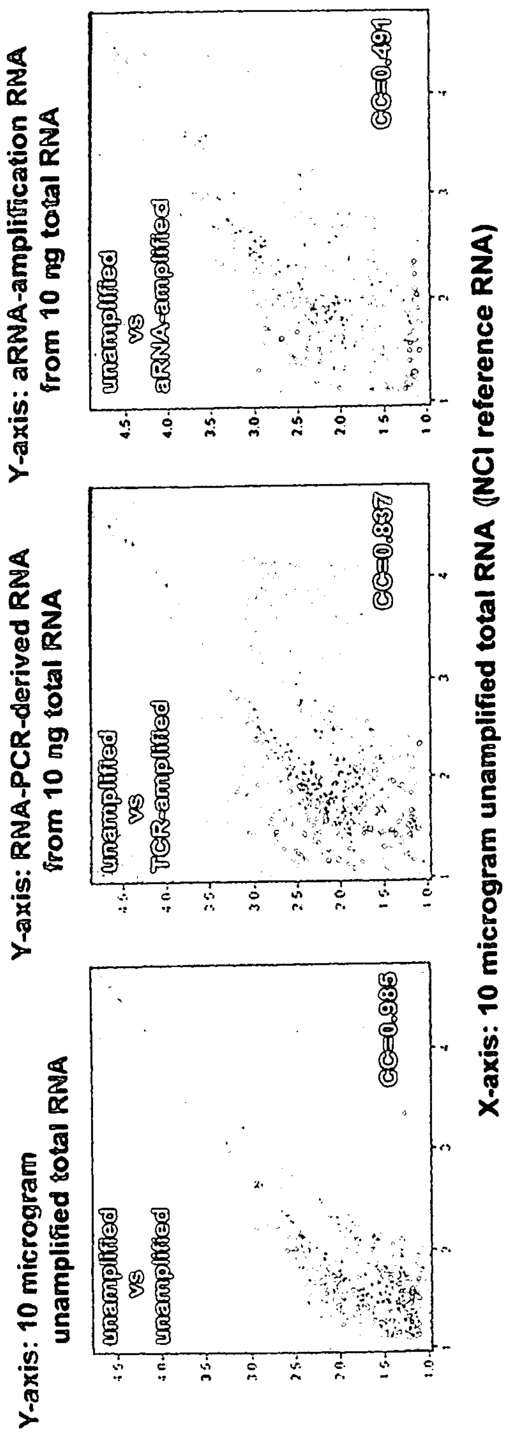


Fig. 4. Microarray analysis using human DNA gene-chips (n = 2 for each group), two-cycle amplification products of RNA-PCR-derived RNA from 10 ng total RNA referenced by National Cancer Institute (NCI) display an average 83.7% correlation coefficiency (CC) compared to 10 µg of the original reference RNA. Because our preset threshold for acceptable variation is onefold change, such high CC rate indicates that >83% of the original mRNA population has been well preserved in almost the same composition and ratio. Traditional aRNA amplification products from 10 ng reference RNA, however, display a lower 49% CC rate, which may result from the use of random primers and, therefore, loss of full-length RNA composition during cycling amplification.

2.3. cDNA Double-Stranding Using DNA Polymerization with Oligo(dG)-Promoter Primers

1. Oligo-$(dG)_{10}$N-T7 RNA promoter primer mix: dephosphorylated 5′-GGCA(GAAT TGTAATACGA CTCACTCACT ATAGGGAAGG CGGGGGGGN (N = A, T, or C; total 100 pmol/µL including 35 pmol/µL 5′-GGCAGTGAAT TC ATACGA CTCACTCACT ATAGGGAAGG CGGGGGGGA-3′, 35 pmol, 5′-GGCAGTGAAT TGTAATACGA CTCACTCACT ATAGGGAAGG CG GGGGT-3′, and 30 pmol/µL 5′-GGCAGTGAAT TGTAATACGA CTCA CACT ATAGGGAAGG CGGGGGGGC-3′).
2. 10X cDNA double-stranding buffer: 500 m*M* Tris-HCl (pH 9.2) at 25°C, 160 r $(NH_4)_2SO_4$, 20 m*M* $MgCl_2$; prepare fresh.
3. cDNA double-stranding reaction mix: 10 µL DEPC-treated ddH_2O, 5 µL of 1 cDNA double-stranding buffer, 2 µL of 10 m*M* dNTP mix (10 m*M* each of dA dGTP, dCTP, and dTTP), 0.7 µL *Taq* DNA polymerase (5 U/µL), and 0.3 Pwo DNA polymerase (5 U/µL); prepare just before use.
4. Incubation chamber: 94°C, 50°C, and 68°C.
5. Purification spin column: 100 bp cutoff filter.

2.4. Generation of Full-Length Sense RNA Using In Vitro Transcription with Promoter-Driven RNA Polymerase

1. 10X In vitro transcription (IVT) buffer: 400 m*M* Tris-HCl (pH 8.0) at 25°C, 100 r $MgCl_2$, 50 m*M* DTT, and 5 mg/mL nuclease-free bovine serum albumin (BSA
2. T7 RNA polymerase (80 U/µL).
3. IVT reaction mix: 8 µL DEPC-treated ddH_2O, 4 µL 10X IVT buffer, 4 µL 10 m*M* dNTP mix (10 m*M* each of ATP, GTP, CTP, and UTP), 2 µL RNa (25 U/µL), and 2 µL T7 RNA polymerase; prepare just before use.
4. Incubation mixer: 37°C, 100*g* vortex for 30 s between every 30-min interval.

2.5. Another Round of Thermocycling Amplification Using RNA-PCR with Sense RNA

1. $Poly(dT)_{24}$ primer: dephosphorylated 5′-dTTTTTTTTTT TTTTTTTTTT TT-3′ (100 pmol/µL).
2. 10X RT&T buffer: 600 m*M* Tris-HCl (pH 8.3) at 25°C, 300 m*M* KCl, 80 r $MgCl_2$, 100 m*M* DTT, and 5*M* betaine.
3. RNA-PCR reaction mix: 2 µL DEPC-treated ddH_2O, 2 µL of 10X RT&T buf 2 µL of 10 m*M* dNTP mix (10 mM each for dATP, dGTP, dCTP, and dTTP), 1 RNasin (25 U/µL), and 2 µL AMV reverse transcriptase (50 U/µL): prep just before use.
4. DEPC-treated ddH_2O.
5. Oligo-$(dG)_{10}$N-T7 RNA promoter primer mix: dephosphorylated 5′-GGCA GAAT TGTAATACGA CTCACTCACT ATAGGGAAGG CGGGGGGGN (N = A, T or C; total 100 pmol/µL).
6. *Taq* DNA polymerase (5 U/µL) and Pwo DNA polymerase (5 U/µL).

7. Incubation chamber: 65°C, 42°C, 50°C, 94°C, and 68°C.
8. Purification spin column: 100 bp cutoff filter.

3. Methods

3.1. Generation of First-Strand cDNA Using Reverse Transcription with Poly(dT) Primers

The starting material can be either 0.1 ng to 1 μg total RNA *(3)* or permeabilized cell preparation *(3,12)* (*see* **Notes 2** and **3**). Poly(A^+) RNA is selected using poly(dT) primers, which contains about 20–26 deoxythymidylate oligonucleotides. The first-strand cDNA is synthesized by reverse transcription from the poly(A^+) RNA with the poly(dT) primers. As shown in **Fig. 1**, the promoter used here is a T7 bacteriophage RNA promoter element.

1. Primer annealing: Suspend RNA in 5 μL of DEPC-treated water, mix well with 1 μL poly$(dT)_{24}$ primer, heat to 65°C for 5 min for minimizing secondary structure, cool to 50°C for 1 min for primer hybridization, and then cool on ice.
2. First-strand cDNA synthesis: Add 14 μL of first reverse transcriptase mix and heat to 42°C for 50 min. Add another 1 μL of reverse transcriptase and mix. Continue to incubate the reaction at 42°C for 30 min, heat to 50°C for 10 min, and then cool on ice. The RNA is still attached noncovalently to the cDNA.
3. Denaturation: Heat the reaction at 94°C for 3 min and then cool on ice immediately.
4. Primer removal and buffer exchange: Load the reaction into a purification spin column, spin for 10 min at 14,000*g*, and discard the flowthrough (*see* **Note 4**). Add 200 μL of DEPC-treated ddH_2O into the spin column to wash the cDNA, spin for 10 min at 14,000*g*, and discard the flowthrough. Add 20 μL of DEPC-treated ddH_2O into the spin column to dissolve the cDNA, place the spin column upside down in a new collecting microtube, and spin 3 min at 3000*g*. Store the 20 μL of the purified cDNA in a –20°C freezer or perform the next step immediately.

3.2. cDNA Amplification Using Terminal Transferase Tailing with Poly(dC) Oligonucleotides

In this method, which was reported by Lin and Ying in 2000, the first-strand cDNA is dC-tailed using TdT and a promoter-linked oligo(dG) primer is applied to initiate the second-strand cDNA synthesis. As shown in **Fig. 1**, the promoter used here is a T7 bacteriophage RNA promoter element. Although this external priming procedure preserves better full-length conformation of the mRNA, the efficiency of the TdT tailing reaction seems to depend on the particular 3′ termini of different first-strand cDNA species, resulting in uneven coverage. Such a problem, however, can be improved by adequate TdT activity rate and constant reaction vortex (*see* **Note 5**). Practically, 1 U of TdT

is required for tailing every picomole of cDNA in a mild shaking incub (100*g*), if the average size of cDNA is 3 kb.

1. TdT tailing reaction: Add 10 μL of terminal transferase reaction mix tc purified cDNA and mix well. Incubate the reaction at 37°C for 30 min occasionally mix the reaction every 5 min for better tailing coverage.
2. Reaction stop: Heat the reaction at 94°C for 3 min and cool on ice immedia

3.3. cDNA Double-Stranding Using DNA Polymerization with Oligo-(dG)-Promoter Primers

The mRNA of the resulting mRNA–cDNA hybrid is denatured, and dou stranded cDNA was formed using a modified two-cycle PCR-like reac with promoter-linked oligo-(dG) primers. Then, the amplification of cD representative can be achieved by in vitro transcription using the promc linked double-stranded cDNA as the template. The full-length construc the cDNA template is protected and flanked by poly(dC)- and poly(d oligonucleotide in its 5′ and 3′ termini, respectively.

1. Primer annealing: Add 2 μL of oligo-$(dG)_{10}$N-T7 RNA promoter primer mi the reaction, heat to 94°C for 3 min for mRNA removal, cool to 50°C for 10 for primer hybridization, and then cool on ice.
2. cDNA double-stranding: Add 20 μL of cDNA double-stranding reaction mi the reaction, mix well, and then incubate the reaction at 68°C for 10 min. Re the thermocycling incubation from 94°C for 3 min, 50°C for 10 min, and t 68°C for 10 min, one more time.
3. Breaking cell membrane: For using permeabilized cells as starting material, 200 μL of 2% nonionic detergent (octylphenoxy)polyethanol to the reaction vortex at 100*g* for 10 min.
4. Primer removal and buffer exchange: Load the reaction into a purification s column, spin for 10 min at 14,000*g*, and discard the flowthrough. Add 200 μL DEPC-treated ddH_2O into the spin column to wash the double-stranded cD spin for 10 min at 14,000*g*, and discard the flowthrough. Add 20 μL of DE treated ddH_2O into the spin column to dissolve the double-stranded cDNA, pl the spin column upside down in a new collecting microtube, and spin for 3 r at 3000*g*. Store the 20 μL of the purified cDNA in a –20°C freezer or perfc the next step immediately.

3.4. Generation of Full-Length Sense RNA Using In Vitro Transcription with Promoter-Driven RNA Polymerase

The promoter of the double-stranded cDNA is now served as a recc nition site for RNA polymerase during an in vitro transcription reacti The in vitro transcription provides linear amplification up to 2000-fold

the amount of starting materials (*6,7*). The proofreading capability of the RNA polymerase ensures the fidelity of the resulting nucleic acid products. Because the promoter is incorporated in the same orientation of mRNA, the resulting product is sense RNA (mRNA) rather than antisense RNA. A cap structure can be added to the sense RNA for further peptide synthesis (*see* **Note 6**).

1. In vitro transcription reaction: Add 20 μL of IVT reaction mix to the purified cDNA and mix well. Incubate the reaction at 37°C for 2–3 h and occasionally mix the reaction every 30 min for better RNA elongation (*see* **Note 7**).
2. Buffer exchange and sample concentration: Load the reaction into a purification spin column, spin for 10 min at 14,000*g*, and discard the flowthrough. Add 200 μL of DEPC-treated ddH_2O into the spin column to wash the poly(A^+) RNA, spin for 10 min at 14,000*g*, and discard the flowthrough. Add 20 μL of DEPC-treated ddH_2O into the spin column to dissolve the poly(A^+) RNA, place the spin column upside down in a new collecting microtube, and spin for 3 min at 3000*g*. Store the 20 μL of the purified poly(A^+) RNA in a –80°C freezer or perform the next step immediately.

3.5. Another Round of Thermocycling Amplification Using RNA-PCR with Sense RNA

The sense RNA so generated is flanked by poly(dC)- and poly(A)-oligonucleotide in its 5′ and 3′ termini, respectively. These homopolymeric tails not only maintain the full-length conformation of the mRNA but also serve as templates for the poly(dT) and promoter-linked oligo-(dG) primers. The cycling of the above transcriptional amplification can be reiterated using the sense RNA directly in the next round of RNA-PCR reaction following the cycling steps of **Subheading 3.4.** and **3.5.** (*see* **Note 8**).

1. Primer annealing: Add 1 μL poly$(dT)_{24}$ primer to 10 μL of the purified poly(A^+) RNA, heat to 65°C for 5 min for minimizing secondary structure, cool to 50°C for 1 min for primer hybridization, and then cool on ice.
2. First-strand cDNA synthesis: Add 9 μL of RNA-PCR reaction mix to the reaction, and heat to 42°C for 50 min. Add another 1 μL of reverse transcriptase and mix. Continue to incubate the reaction at 42°C for 30 min, heat to 50°C for 10 min, and then cool on ice. The RNA is still attached noncovalently to the cDNA.
3. Denaturation: Add 16 μL of DEPC-treated ddH_2O and 2 μL of oligo-$(dG)_{10}$N-T7 RNA promoter primer mix to the reaction and incubate the reaction at 94°C for 3 min and then 50°C for 10 min.
4. cDNA double-stranding: Add 0.7 μL *Taq* DNA polymerase and 0.3 μL Pwo DNA polymerase to the reaction and incubate at 68°C for 10 min.
5. Primer removal and buffer exchange: Load the reaction into a purification spin column, spin for 10 min at 14,000*g*, and discard the flowthrough. Add 200 μL of

DEPC-treated ddH_2O into the spin column to wash the double-stranded cD spin for 10 min at 14,000*g*, and discard the flowthrough. Add 20 μL of DE treated ddH_2O into the spin column to dissolve the double-stranded cDNA, pl the spin column upside down in a new collecting microtube, and spin for 3 n at 3000*g*. Store the 20 μL of the purified cDNA in a –20°C freezer or perfo the next step immediately.

6. cDNA library assessment (*see* Chapter 13).

4. Notes

1. When performed with a specific primer complementary to the 3′-targe sequence of a desired mRNA in conjunction with the oligo-(dG)-T7 primer, th end of a specific mRNA can be generated by RNA-PCR for 5′-UTR (untransla region) analysis. The design of these sequence-specific primers is based on same principle used by PCR (50–55% G-C rich). On the other hand, in additi to the 3′-end sequence-specific primer, we can also use another sequence-speci promoter–primer to amplify certain domain within the code-reading fra of the mRNA for further research. The design of these promoter–prime however, requires a higher G-C content (60–65%) working at the same anne ing temperature as the sequence-specific primers because of their unmatch promoter regions. For example, the new annealing temperature for the sequen matched region of a promoter-primer is [2(dA + dT) + 3(dC + dG)] (5/6), including the promoter region. Please remember that all primers were purified polyacrylamide gel electrophoresis before used in an RNA-PCR reaction.
2. Isolated cells were preserved in 500 μL of ice-cold 10% formaldehyde in susp sion buffer (0.15 *M* NaCl [pH 7.0], 1 m*M* EDTA) for the following fixation a permeabilization procedure (*12*). After 1-h incubation with occasional agitati fixed cells were collected with a Microcon-50 filter and washed by 350 μL ice-cold PBS with vigorous pipetting. The collection and wash were repeated least once. The fixed cells were then permeabilized in 500 μL of 0.5% nonio detergent (octylphenoxy)polyethanol for 1 h with frequent agitation. After th three collections and washes were given to cells, as earlier, but using 350 μL ice-cold phosphate-buffered saline (PBS) containing 0.1 *M* glycine inste The cells were finally mixed with 0.1 μ*M* poly$(dT)_{24}$ primer and resuspend in the same buffer with vigorous pipetting to evenly distribute them into sm aliquots (about 50 cells in 10 μL) for RNA-PCR. They could be stored at –80 for up to 2 wk.
3. The adequate amount of fixed cells for RNA-PCR ranged from more than but less than 200 cells (if each cell contains 0.1 pg mRNAs) because of t completeness of TdT tailing reactions. The chance to generate a good mR library (300 bp~ to 5 kb) from <20 cells is less than 50% based on our tes The chance to generate a complete mRNA library from more than 200 ce will depend on the relative amount of terminal transferase (TdT) activity to t first-strand cDNA molecules. The TdT activity is less effective when too ma cDNA molecules interact with TdT in a limited tailing reaction. In brief,

currently know that the concentration of TdT determines the completeness of an RNA-PCR-derived library, whereas that of RNA polymerases and reverse transcriptases determines the amplification rate of a RNA-PCR reaction. Therefore, we suggest that please use at least 50 U of TdT for every 0.1 ng mRNAs in a tailing reaction and more than 60 U each of RNA polymerases and reverse transcriptases in a 20-μL transcription reaction.

4. Relative Centrifugal Force (RCF) (g) = $(1.12 \times 10^{-5}) \cdot (\text{rpm})^2 \cdot r$, where r is the radius in centimeters measured from the center of the rotor to the middle of the spin column and rpm is the speed of the rotor in revolutions per minute.
5. The first-strand cDNA is poly(dC)-tailed by TdT using the provided condition that should produces an average 8–15-base overhang. The incorporation rate is increased about 75–85% with occasionally gentle mix, but drops to about 50% without mixing. The efficiency of TdT tailing seems to be varied among different mRNA species but can be improved by occasionally gently mixing in a short period of incubation time. The length of homopolymeric tails should be limited by the special designs of returning primers or promoter–primers, as mentioned in **ref.** *(3)* (e.g., an equal mixture of T7-oligo-$(dG)_{10-12}$N primers; N= dA, dT, or dC). The homopolymeric region of a returning primer should range from 7 to 16 bases, most preferably from 10 to 12 bases. The use of cobalt-based buffers is not recommended in this protocol.
6. The RNA-PCR has been tested to provide amplified full-length mRNAs for in vitro translation (*see* Chapter 25). A cap nucleotide can be added to the 5′ end of the amplified mRNAs during the transcriptional amplification. Unlike normalized RNAs, the capped mRNAs can be directly used in protein synthesis and may help to isolate such protein activity if its folding is correct. The preferred cap nucleotides include P1-5′-(7-methyl)-guanosine-P3-5′-adenosine-triphosphate and P1-5′-(7-methyl)-guanosine-P3-5′-guanosine-triphosphate. Such protein products are useful for protein differential display on a two-dimensional gel.
7. The most stable and efficient IVT reaction occurs during the first 2-h of incubation at 37°C. The rate of RNA synthesis decreases considerably (40–50%) after a 3-h incubation or below 37°C incubation. A longer reaction may increase yield, but the possibility of degradation by RNase increases. Occasionally, gentle mixing can prevent the stall of crowded RNA polymerases on a template and enhance full-length synthesis. The overall rate of RNA polymerization is maximal between pH 7.7 and 8.3, but it remains about 70% of maximum at pH 7.0 or 9.0. High concentrations of NaCl, KCl, or NH_4Cl above 75 m*M* will inhibit the reaction.
8. To reach about 1.5×10^7-fold amplification of mRNAs, 3–4 cycles of RNA-PCR were needed to perform for about 20–50 cells and 2–3 cycles for about 125–200 cells. The optical density (OD) (A_{260}/A_{280}) value ranged from about 1.7 to 2.0 for mRNA products and mRNA–cDNA hybrids and from about 1.6 to 1.9 for double-stranded cDNA products, depending at which cycling step you stop the RNA-PCR reaction. Remove enzymes with protein-remover filters (Microcon) before OD detection. A lower OD value may indicate an insufficient amplification rate, enzyme-related variation, or RNase contamination. preferred to run a 1% formaldehyde–agarose gel to exam the quality of RN PCR products. Identification of some rare mRNAs (<six copies/cell) by RT-P from the RNAPCR-derived library is another way to observe its quality.

References

1. Thykjaer, T., Workman, C., Kruhoffer, M., Demtroder, K., Wolf, H., Anders L. D., et al. (2001) Identification of gene expression patterns in superficial ε invasive human bladder cancer. *Cancer Res.* **61,** 2492–2499.
2. Crino, P. B., Trojanowski, J. Q., Dichter, M. A., and Eberwine, J. (2001) Embryo neuronal markers in tuberous sclerosis: single-cell molecular pathology. *Pr Natl. Acad. Sci. USA* **93,** 14,152–14,157.
3. Lin, S. L., Chuong, C. M., Widelitz, R. B., and Ying, S. Y. (1999) In vivo analy of cancerous gene expression by RNA-polymerase chain reaction. *Nucleic A Res.* **27,** 4585–4589.
4. Goh, S. H., Park, J. H., Lee, Y. J., Lee, H. G., Yoo, H. S., Lee, I. C., et al. (20(Gene expression profile and identification of differentially expressed transcri during human intrathymic T-cell development by cDNA sequencing analys *Genomics* **70,** 1–18.
5. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Construction and analysis cDNA libraries, in *Molecular Cloning*, 2nd ed. Cold Spring Harbor Laborato Cold Spring Harbor, NY, pp. 8.11–8.19.
6. Ying, S. Y., Liu, H. M., Lin, S. L., and Chuong, C. M. (1990) Amplified RN synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Ac Sci. USA* **87,** 1663–1667.
7. Eberwine, J., Yeh, H., Miyashrio, K., Cao, Y., Nair, S., Finnell, R., et al. (199 Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci. U* **89,** 3010–3014.
8. Lin, S. L., Chuong, C. M., and Ying, S. Y. (2001) A novel mRNA–cDNA interf ence phenomenon for silencing bcl-2 expression in human LNCaP cells. *Bioche Biophys. Res. Commun.* **281,** 639–644.
9. Compton, J. (1991) Nucleic acid sequence-based amplification. *Nature* **35** 91, 92.
10. O'Dell, D. M., Raghupathi, R., Crino, P. B., Morrison, B., 3rd, Eberwine, J. I and McIntosh, T. K. (1998) Amplification of mRNAs from single, fixed, TUNE positive cells. *BioTechniques* **25,** 566–570.
11. Wang, E., Miller, L. D., Ohnmacht, G. A., Liu, E. T., and Marincola, F. M. (200 High-fidelity mRNA amplification for gene profiling. *Nat. Biotech.* **18,** 457–45
12. Embleton, M. J., Gorochov, G., Jones, P. T., and Winter, G. (1992). In-cell PC from mRNA: amplifying and linking the rearranged immunoglobulin heavy a light chain V-genes within single cells. *Nucleic Acids Res.* **20,** 3831–3837.

16

Single-Cell mRNA Library Analysis by Northern Blot Hybridization

Shi-Lung Lin

1. Introduction

The debut of RNA-polymerase cycling reaction (RNA-PCR) has promised to provide linear amplification of a reproducible mRNA library from as few as 20 single cells *(1)*. By incorporating a RNA promoter element during the synthesis of double-stranded complementary DNA (cDNA) templates, a poly(A^+) RNA library can be generated and reamplified from the templates in the same conformation and composition as its mRNA origins (Fig. 1). Using microarray analysis, the RNA-PCR-derived poly(A^+) library has been proven to contain above 97% of the original poly(A^+) RNA population and maintain $88 \pm 4\%$ linear correlationship to the populationary ratio of each RNA species (Chapter 12). It has also been tested to generate a full-length mRNA library from as few as 20 homologous tissue cells (2-pg mRNAs) for profiling cancer stages in vivo.

Northern blot analysis of gene expression usually requires abundant mRNA resources (>0.5 μg/lane), which is impossible to acquire from a few homologous tissue cells using traditional RNA extraction methods. However, we have acquired 30 μg of amplified poly(A^+) RNAs in one 50-μL reaction after three rounds of RNA-PCR amplification from about 20 single cells. This represents a 1.5×10^6-fold increase based on the comparison between the amount of the amplified poly(A^+) RNAs and that of theoretically presumed mRNAs within a cell (0.1 pg). It is noted that some rare RNAs can be well preserved by RNA-PCR for further gene analysis. Therefore, RNA-PCR can be a tool for providing unlimited mRNA resources for Northern blot detection at the single-cell scale.

From: *Methods in Molecular Biology, vol. 221: Generation of cDNA Libraries: Methods and Protocols*
Edited by: S.-Y. Ying © Humana Press Inc., Totowa, NJ

25

Peptide Library Construction from RNA-PCR-Derived RNAs

Shi-Lung Lin

1. Introduction

The generation of peptide from messenger RNA (mRNA) provides a convenient source for current proteomic analysis. Intron-free mRNA possessing adenine-uracil-guanine (AUG) start codons can be translated into labeled or unlabeled peptides under a predetermined reticulocyte lysate condition. In conjunction with RNA-polymerase cycling reaction (*see* **Fig. 1**; RNA-PCR), full-length gene transcripts can be unlimitedly amplified for protein/peptide synthesis in vitro *(1)*. Many commercialized in vitro translation systems provide a cap nucleotide, which can be added to the 5′ end of the amplified poly(A^+) RNAs during the transcription step of RNA-PCR. Totally resembling mRNAs, the capped poly(A^+) RNAs can be used to synthesize proteins/peptides with labeling and may help the functional analysis of protein activity if they fold correctly.

We has successfully tested the RNAPCR-derived protein analysis in a prostatic cancer cell line, LNCaP, of which the protein data matched previous findings using mRNA. The antiapoptotic gene family of *bcl-2* has been well known for their ability to increase cancer resistance to multiple anticancer drugs. Previous data has predicted that a mutated form of *bcl-2* may be elevated in the drug-resistant LNCaP cells after androgen retrieval treatment *(2)*. Using Northern and Western blots, we confirmed that a truncated form of the *bcl*-2 member can be clearly detected in both the mRNA and protein levels, indicating a consistent result between these two methods. Because the folding of synthesized proteins may be different from that of the original one, the detection of such mutated changes will depend on the antibody used. For the

From: *Methods in Molecular Biology, vol. 221: Generation of cDNA Libraries: Methods and Protocols*
Edited by: S.-Y. Ying © Humana Press Inc., Totowa, NJ